Embedded data signaling

The invention relates to signaling of embedded data, and also to the formatting and handling of data streams with embedded data.

5   The syntax definition of several audio coding schemes/standards (e.g. mp3 and MPEG-AAC, see e.g. ISO/IEC 13818-3 and 13818-7 respectively) provides the possibility of adding ancillary/embedded data to coded audio streams. Compliant decoders are only required to parse the embedded data, not to interpret it. In practice, the embedded data is often used to store a coded data stream related to an enhancement coder/tool (e.g. mp3PRO,

10   MPEG-4 AAC+SBR, wherein "SBR" stands for Spectral Band Replication). Such an enhanced coder/tool can be used on top of the core coder to improve the quality of the core audio stream. Since a non-enhanced coder is required to parse the embedded data, the data embedding is done in backward compatible manner.

        In MPEG-4 (see ISO/IEC 14496-1 for MPEG-4 Systems and ISO/IEC 14496-

15   3 for MPEG-4 Audio), signaling of stream content is done by means of descriptors. Each elementary stream (i.e. a consecutive flow of mono-media data such as audio or video which can be packetized) has a corresponding descriptor. The current descriptor definition does not provide for signaling of embedded data. Signaling of the embedded data can of course be realized by means of a corrigendum on the descriptors. However, such a corrigendum cannot

20   be implemented such that the standard remains backward compatible with the current definition. Alternatively, one could use a descriptor in the embedded data itself. This has the disadvantage that the embedded data is not signaled at elementary stream level and that therefore the embedded data needs to be accessed to see what it contains.

25

        An object of the invention is to provide advantageous signaling of embedded data. To this end, the invention provides a method, an encoder, a signal, a storage medium, a method of decoding, a decoder, a transmitter or recorder and a receiver as defined in the independent claims. Advantageous embodiments are defined in the dependent claims.

According to a first aspect of the invention, a signal representing main data is provided, the main data including embedded data, the main data being provided with a main data descriptor for signaling content included in the main data, wherein an embedded data descriptor is formed for signaling content included in the embedded data, and wherein the

5   embedded data descriptor is provided outside (or separate from) the main data and the main data descriptor. By providing the embedded data descriptor outside the main data descriptor, a corrigendum of a currently defined set of main data descriptors is not necessary to take into account non-defined content included in the embedded data; a signal having new embedded data descriptors remains backward compatible as regards the main data and the main data

10   descriptors. By providing the embedded data descriptor outside the main data and thus also outside the embedded data, the main data remains compatible and it is further not necessary to access the embedded data itself to get a description of it.

As a result, elementary streams with embedded data may have two respective descriptors, one for the main data and another for the embedded data.

15   The main data, the embedded data, the main data descriptor and the embedded data descriptor may all be present in the same transport signal.

Advantageously, the signal is an elementary audio data stream, wherein the embedded data is enhancement data to enhance the audio available in the rest of the elementary data stream. The enhancement data is preferably information suitable for

20   performing spectral band replication. Alternatively, the enhancement data is suitable for extension of the number of channels, e.g. from 1 to 2 channels or from 2 to 5 channels, see e.g. the paper of Faller and Baumgarte, "Binaural cue coding applied to stereo and multi-channel audio compression", *AES 112th paper 5574*, May 10-13, 2002, Germany and e.g. European patent application nr. 02076588.9 filed 22 April 2002.

25   Embodiments of the invention are especially useful in those standards in which it is possible to implement a separate embedded data descriptor such that a conventional coder will not wrongfully use its contents, e.g. by ignoring the embedded data descriptor e.g. simply because it uses a non-recognizable code which orders a decoder to ignore the information. Examples of standards where this is easily possible are MPEG-4

30   systems and RFC 3016.

These and other aspects of the invention will be apparent from the elucidated with reference to the accompanying drawings.

In the drawings:

Fig. 1 shows an example of a conventional elementary stream;

Fig. 2 shows an example of an elementary stream provided with an embedded data descriptor according to an embodiment of the invention;

5   Fig. 3 shows a system according to an embodiment of the invention.

The drawings only show those elements that are necessary to understand the embodiments of the invention.

10   Fig. 1 shows an example of a conventional elementary stream ES. The elementary stream ES may be a packetized elementary stream. The elementary stream ES comprises main data MD and a main data descriptor MDD. An exemplary descriptor MDD for an encoded audio stream may be as follows:

15   MDD

{

Audio object type ("AOT")

Sampling frequency

Channel configuration

20   AOT specific configuration information

}

As an example, configuration information specific to AAC related AOT's include a frame length, i.e. the number of PCM samples per channel related to one AAC

25   audio frame.

Further, the main data MD includes embedded data ED. The main data MD preferably comprises encoded audio data, e.g. AAC or mp3 encoded data. It is also possible that the main data MD comprises video data. The embedded data ED preferably includes enhancement data to enhance the main data MD, e.g. by spectral band replication in the case

30   of audio or by spatial, SNR or other enhancement for video. Alternatively, the enhancement data is suitable for extension of the number of channels, e.g. from 1 to 2 channels or from 2 to 5 channels as indicated above.

In some systems, e.g. in MPEG-4, the data descriptor MDD is not concatenated with the main data MD in the elementary stream, but is provided separately. To

determine which descriptor relates to which elementary stream, some identification is used in both the descriptor as well as the elementary stream ES.

The embedded data ED is parsed in a decoder and recognized by an enhanced decoder which is able to use the enhancement data present in ED. Usually, the embedded

5   data ED includes some kind of identification/description to make identification of the enhancement data ED possible, although in proprietary systems it is also possible to agree between an encoder and a decoder that the embedded data ED always comprises enhancement data according to a predetermined format.

Fig. 2 shows an example of a further elementary stream EES provided with an

10   embedded data descriptor EDD according to an embodiment of the invention. The embedded data descriptor EDD includes identification information to make identification of the type of embedded data ED possible. The descriptor EDD may also include other useful information. An exemplary descriptor EDD for the data embedded in an encoded audio stream may be as follows:

15

EDD

{

            Audio (enhancement) object type ("AOT")
            AOT specific configuration information

20   }

The definition of the EDD strongly depends on the audio (enhancement) object type. In the case of SBR, it contains the sampling frequency mode, which can be single- or multi-rate. In the case of channel extension, the embedded data descriptor may

25   contain information on the extended channel configuration.

The embedded data descriptor EDD is provided outside the main data MD and the main data descriptor MDD and is therefore easily accessible. Depending on the coding scheme used the data descriptors MDD and EDD may be supplied in a concatenated way with the main data MD. It is also possible to provide the descriptors separately in another part

30   of the signal, e.g. all descriptors grouped together. Some linking information is then necessary to relate the descriptors to the relevant elementary streams.

MPEG-4 implementation

The above described embodiment of the invention is advantageously applied in an MPEG-4 or MPEG-4 like coding scheme. In MPEG-4, the main data descriptor MDD and the embedded data descriptor EDD are provided separately with respect to the

5　elementary stream EES. MPEG-4 systems provides tools to relate the descriptors to the relevant elementary streams.

RFC3016 implementation

In RFC 3016 (IETF RFC 3016: "RTP Payload Format for MPEG-4

10　Audio/Visual Streams", Kikuchi Y. et al., November 2000), descriptor information is provided as a Session Description Protocol (SDP) parameter. For example in case of audio the audio decoding configuration is described by the parameter "config" as a hexadecimal string that represents the audio descriptor as defined by MPEG-4. An example is given below:

15　config=000001B001000001B509000001000000120008440FA282C2090A21F.
Another descriptor can be added by defining a new parameter, such as embedded-data-config. Receivers are required to ignore new or unknown parameters.

System according to an embodiment of the invention

20　Fig. 3 shows a system according to an embodiment of the invention. The system comprises an apparatus 1 for transmitting or recording an encoded signal [S]. The apparatus 1 comprises an input unit 10 for obtaining an input signal S, e.g. an audio and/or video signal. The input unit 10 may be an antenna, microphone, network connection, etc. The apparatus 1 further comprises an encoder 11 for encoding the signal S according to an above

25　described embodiment of the invention (see in particular Fig. 2) in order to obtain an encoded signal comprising main data MD including embedded data ED, and the descriptors MDD and EDD. The encoded signal is furnished to an output unit 12 which formats the main data MD including the embedded data ED, and the descriptors MDD and EDD into an encoded signal [S] having a suitable format for transmission or storage via a transmission medium or storage

30　medium 2 (e.g. as defined in RFC 3016). The system further comprises a receiver or reproduction apparatus 3 which receives the encoded signal [S] in an input unit 30. The input unit 30 furnishes the main data MD, the embedded data ED and the data descriptors MDD and EDD to the decoder 31. The decoder 31 decodes the encoded signal by performing a decoding process which is substantially an inverse operation of the encoding in the encoder

11 wherein a decoded signal S' is obtained which corresponds to the original signal S except for those parts which were lost during the encoding process. The decoder 31 furnishes the decoded signal S' to a reproduction unit 32 such as a speaker for reproducing the decoded signal S'. The reproduction unit 32 may also be a transmitter for further transmitting the

5   decoded signal S' for example over an in-home network, etc.

Existing receivers are able to ignore the EDD as described above for the RFC 3016 case. Future receiver implementations may be able to interpret the EDD. In this case passing of the EDD to unit 31 may be dependent on the capabilities of unit 31. For example, in those implementations wherein the decoder 31 does not support a feature to which the

10   embedded data relates, then the input unit 30 may decide not to provide the EDD to unit 31 in order to save bandwidth.

Embodiments of the invention may be applied in audio and/or video broadcast, Internet Radio, 3GPP, Internet distribution, Solid State Audio, 3G terminals, GPRS and commercial successors thereof.

15   It should be noted that the above-mentioned embodiments illustrate rather than limit the invention, and that those skilled in the art will be able to design many alternative embodiments without departing from the scope of the appended claims. In the claims, any reference signs placed between parentheses shall not be construed as limiting the claim. This word 'comprising' does not exclude the presence of other elements or steps than those listed

20   in a claim. The invention can be implemented by means of hardware comprising several distinct elements, and by means of a suitably programmed computer. In a device claim enumerating several means, several of these means can be embodied by one and the same item of hardware. The mere fact that certain measures are recited in mutually different dependent claims does not indicate that a combination of these measures cannot be used to

25   advantage.